

CUBE SPACES AND THE MULTIPLE TERM RETURN TIMES THEOREM

PAVEL ZORIN-KRANICH

ABSTRACT. We give a new proof of Rudolph's multiple term return times theorem based on Host-Kra structure theory. Our approach also yields a multiple term Wiener-Wintner-type return times theorem for nilsequences.

1. INTRODUCTION

In this article we are concerned with universally good weights for pointwise convergence of ergodic averages, i.e., sequences (a_n) such that, for every measure-preserving system (Y, S) and every $g \in L^\infty(Y)$, the averages

$$\lim_N \frac{1}{N} \sum_{n=1}^N a_n g(S^n y)$$

converge for a.e. $y \in Y$. Bourgain's return times theorem [BFKO89] asserts that, given any ergodic measure-preserving system (X, T) , for every $f \in L^\infty(X)$ and a.e. $x \in X$ the sequence of weights $a_n = f(T^n x)$ is universally good for pointwise convergence. The name "return times theorem" comes from the case of a characteristic function $f = 1_A$, $A \subset X$. Then the theorem can be equivalently formulated by saying that, for a.e. $x \in X$, the pointwise ergodic theorem on any system Y holds along the sequence of return times of x to A .

This has been extended to averages involving multiple terms by Rudolph [Rud98]. In order to formulate his result and for future convenience we now introduce some notation. By a *system* we mean an ergodic regular measure-preserving system (X, μ, T) with a distinguished countable subset $D \subset L^\infty(X)$ that is sufficiently large (we will formulate the precise condition on D in Definition 2.8).

Definition 1.1. Let P be a statement about ergodic regular measure-preserving systems (X_i, μ_i, T_i) , functions $f_i \in L^\infty(X_i)$ and points $x_i \in X_i$, $i = 0, \dots, k$. We say that P holds for *universally almost every* (u.a.e.) tuple x_0, \dots, x_k if

- (0) For every system (X_0, μ_0, T_0, D_0) there exists a set of full measure $\tilde{X}_0 \subset X_0$ such that
- (1) for every system (X_1, μ_1, T_1, D_1) there exists a measurable set $\tilde{X}_1 \subset X_0 \times X_1$ such that for every $\vec{x}_0 \in \tilde{X}_0$ the set $\{x_1 : (\vec{x}_0, x_1) \in \tilde{X}_1\}$ has full measure in X_1 and
- \vdots
- (k) for every system (X_k, μ_k, T_k, D_k) there exists a measurable set $\tilde{X}_k \subset X_0 \times \dots \times X_k$ such that for every $\vec{x}_{k-1} \in \tilde{X}_{k-1}$ the set $\{x_k : (\vec{x}_{k-1}, x_k) \in \tilde{X}_k\}$ has full measure in X_k and

we have $P(f_0, \dots, f_k, \vec{x}_k)$ for every $\vec{x}_k \in \tilde{X}_k$ and every $f_i \in D_i$, $i = 0, \dots, k$.

Date: October 19, 2012.

2010 Mathematics Subject Classification. 28D05.

Key words and phrases. cube space, return times theorem, Wiener-Wintner theorem.

With this convention Rudolph's multiple term return times theorem says that for every $k \in \mathbb{N}$ the averages

$$(1.2) \quad \frac{1}{N} \sum_{n=1}^N f_0(T_0^n x_0) \cdots f_k(T_k^n x_k)$$

converge for u.a.e. x_0, \dots, x_k . We call this statement RTT(k). Birkhoff's pointwise ergodic theorem [Bir31] is essentially RTT(0) and Bourgain's return times theorem is RTT(1). More about the history of these and related results can be found in a recent survey by Assani and Presser [AP12a].

It is known that the Host-Kra-Ziegler nilfactor $\mathcal{Z}_k(X_0)$ is characteristic for RTT(k) in the sense that if $f_0 \perp \mathcal{Z}_k(X_0)$, then the averages (1.2) converge to zero u.a.e. [AP12b, Theorem 4]. However, the proof of this fact hitherto depends on the convergence result RTT(k).

In this article we prove both results, RTT(k) and characteristicity, simultaneously by induction on k using the Host-Kra structure theory. We also obtain the following Wiener-Wintner return times theorem for nilsequences thereby generalizing [ALR95, Theorem 1].

Theorem 1.3 (Wiener-Wintner return times theorem for nilsequences). *Let $k, l \in \mathbb{N}$ and $f_i \in L^\infty(X_i)$, $i = 0, \dots, k$. Then for u.a.e. x_0, \dots, x_k and every l -step nilsequence $(a_n)_n$ the averages*

$$\frac{1}{N} \sum_{n=1}^N a_n \prod_{i=0}^k f_i(T_i^n x_i)$$

converge (to zero if in addition $f_0 \perp \mathcal{Z}_{k+1}(X_0)$ or $f_i \perp \mathcal{Z}_{k+l+1-i}(X_i)$ for some $i = 1, \dots, k$).

The general strategy is to consider not only the X_i 's but also the Host-Kra cube spaces of all orders simultaneously and to use the convergence criterion due to Bourgain, Furstenberg, Katznelson and Ornstein (Proposition 2.11) to pass from k to $k+1$. After a preparatory Section 2 we formulate our central convergence Theorem 3.2 on cube spaces, generalizing RTT(k) and giving some information about characteristic factors. The remaining part of Section 3 is devoted to the proof of Theorem 3.2. Theorem 1.3 is then proved in Section 4.

2. NOTATION AND TOOLS

Ergodic decomposition. For the purposes of this article we find it illuminating to think of the ergodic decomposition in a particular way (that will be generalized in Section 3). Let (X, μ, T) be a regular ergodic measure-preserving system, i.e. X is a compact metric space, $T : X \rightarrow X$ is an invertible continuous map and μ is a T -invariant ergodic Borel probability measure. By the pointwise ergodic theorem a.e. $x \in X$ is generic for some T -invariant Borel probability measure m_x on X , i.e. $\frac{1}{N} \sum_{n=1}^N f(T^n x) \rightarrow \int f dm_x$ for every $f \in C(X)$. It follows easily that the function $x \mapsto m_x$ is measurable and

$$(2.1) \quad \mu = \int m_x d\mu(x)$$

In particular, for μ -a.e. x the measure m_y is defined for m_x -a.e. y . To see that m_x is ergodic for μ -a.e. x it suffices to verify that

$$(2.2) \quad \int \int \left| \int f dm_y - \int f dm_x \right|^2 dm_x(y) d\mu(x) = 0 \text{ for every } f \in C(X),$$

since this says precisely that the ergodic averages of f converge pointwise m_x -a.e. to an m_x -essentially constant function for μ -a.e. x , and the latter full measure set can be chosen independently from f since $C(X)$ is separable. By definition of m_x, m_y , the dominated convergence theorem and (2.1) we can rewrite the integral in (2.2) as

$$\begin{aligned}
& 2 \lim_N \int \left(\frac{1}{N} \sum_{n=1}^N T^n f \right)^2(x) d\mu(x) - 2 \lim_N \int \left(\frac{1}{N} \sum_{n=1}^N T^n f \right)(x) \int \left(\frac{1}{N} \sum_{n=1}^N T^n f \right)(y) dm_x(y) d\mu(x) \\
&= 2 \lim_N \int \left(\frac{1}{N} \sum_{n=1}^N T^n f \right)^2(x) d\mu(x) - 2 \lim_N \lim_M \int \left(\frac{1}{N} \sum_{n=1}^N T^n f \right)(x) \left(\frac{1}{M} \sum_{m=1}^M T^m f \right)(x) d\mu(x),
\end{aligned}$$

and this vanishes by the pointwise ergodic theorem and the dominated convergence theorem.

Host-Kra structure theory. We recall the basic definitions and main results surrounding the uniformity seminorms [HK05]. Let (X, μ, T) be a regular ergodic measure-preserving system. The cube measures $\mu^{[l]}$ on $X^{[l]} := X^{2^l}$ are defined inductively starting with $\mu^{[0]} := \mu$. In the inductive step, given $\mu^{[l]}$, fix an ergodic decomposition

$$\mu^{[l]} = \int_{X^{[l]}} m_x d\mu^{[l]}(x)$$

as in (2.1). The space on which m_x is defined can be inferred from the subscript x . Define

$$(2.3) \quad \mu^{[l+1]} := \int_{X^{[l]}} \delta_x \otimes m_x d\mu^{[l]}(x).$$

To see that this coincides with the conventional definition one can use (2.1) and (2.2) to write the above integral as

$$\begin{aligned}
(2.4) \quad \mu^{[l+1]} &= \int \int \delta_y \otimes m_y dm_x(y) d\mu^{[l]}(x) \\
&= \int \int \delta_y \otimes m_x dm_x(y) d\mu^{[l]}(x) = \int m_x \otimes m_x d\mu^{[l]}(x).
\end{aligned}$$

The *uniformity seminorms* are

$$(2.5) \quad \|f\|_{U^{l+1}}^{2^{l+1}} := \int \otimes_{\epsilon \in \{0,1\}^{l+1}} f d\mu^{[l+1]} = \int \mathbb{E} \left(\otimes_{\epsilon \in \{0,1\}^l} f | \mathcal{J}^{[l]} \right)^2 d\mu^{[l]},$$

where $\mathcal{J}^{[l]}$ is the $T^{[l]}$ -invariant sub- σ -algebra on $X^{[l]}$. For $f^\epsilon \in L^\infty(X)$, $\epsilon \in \{0,1\}^l$, we will abbreviate $f^{[l]} := \otimes_{\epsilon \in \{0,1\}^l} f^\epsilon$. The uniformity seminorms satisfy the *Cauchy-Schwarz-Gowers inequality* [HK05, Lemma 3.9.(1)]

$$(2.6) \quad \left| \int f^{[l]} d\mu^{[l]} \right| \leq \prod_{\epsilon \in \{0,1\}^l} \|f^\epsilon\|_{U^l}.$$

The U^{l+1} -seminorms determine factors $\mathcal{Z}_l(X)$ by the relation

$$f \perp L^2(\mathcal{Z}_l(X)) \iff \|f\|_{U^{l+1}} = 0 \text{ for } f \in L^\infty(X).$$

The factors $\mathcal{Z}_l(X)$ are inverse limits of rotations on l -step nilmanifolds (in fact, topological inverse limits [HKM10, Theorem A.1]).

We also need the classical fact that the Kronecker factor is characteristic for L^2 convergence of ergodic averages with arbitrary bounded scalar weights, see e.g. [HK09, Corollary 7.3] for a more general version.

Lemma 2.7. *Let (X, T) be an ergodic measure-preserving system and $f \in L^2(X)$ be orthogonal to $\mathcal{Z}_1(X)$. Then for any bounded sequence $(a_n)_n$ one has*

$$\lim_N \frac{1}{N} \sum_{n=1}^N a_n T^n f = 0 \quad \text{in } L^2(X).$$

Conventions about cube measures. In the sequel we will have to consider systems for which a certain approximating procedure can be carried out within their distinguished sets.

Definition 2.8. A system is a regular ergodic measure-preserving system (X, μ, T) with a distinguished set $D \subset L^\infty(X)$ that satisfies the following conditions.

- (1) (Cardinality) D is countable.
- (2) (Density) D contains an L^∞ -dense subset of $C(X)$.
- (3) (Algebra) D is closed under pointwise multiplication and absolute value.
- (4) (Decomposition) For every $f \in D$ and $l \in \mathbb{N}$ there exist decompositions

$$\text{Dec}(l) \quad f = f_\perp + f_{\mathcal{Z}_l} + f_{\text{err},j}, \quad j \in \mathbb{N},$$

such that $f_\perp, f_{\mathcal{Z}_l}, f_{\text{err},j} \in D$, $f_\perp \perp \mathcal{Z}_l(X)$, $f_{\mathcal{Z}_l} \in C(Z_j)$, where Z_j is a nilfactor of $\mathcal{Z}_l(X)$, $\|f_{\text{err},j}\|_{L^\infty(\mu)}$ is uniformly bounded in j and $\|f_{\text{err},j}\|_{L^1(\mu)} \rightarrow 0$ as $j \rightarrow \infty$.

Clearly, for any regular ergodic measure-preserving system (X, μ, T) any countable subset of $L^\infty(X)$ is contained in a set D that satisfies the above conditions.

Lemma 2.9. Let (X, μ, T, D) be a system. Then there exist measurable subsets $Y_l \subset X^{[l]}$ such that for every $l \in \mathbb{N}$ the following statements hold.

- (1) $\mu^{[l]}(Y_l) = 1$ and for every $y \in Y_l$ we have $m_y(Y_l) = 1$.
- (2) For every $y \in Y_l$ the measure m_y is ergodic and one has

$$(2.10) \quad m_y \otimes m_y = \int_{Y^{[l+1]}} m_x d(m_y \otimes m_y)(x).$$

- (3) $Y_l \subset (\tilde{X})^{[l]}$, where $\tilde{X} \subset X$ is the set of points that are generic for each $f \in D$ w.r.t. μ .
- (4) For every $y \in Y_l$, every $k \in \mathbb{N}$ and any functions $f_\epsilon \in D$, $\epsilon \in \{0, 1\}^l$, such that $f_\epsilon \perp \mathcal{Z}_{k+l}(X)$ for some ϵ we have $f^{[l]} \perp \mathcal{Z}_k(X^{[l]}, m_y)$.

Proof. The fact that the last condition holds for full measure sets of y_l follows from the Cauchy-Schwarz-Gowers inequality (2.6). The remaining conditions can be achieved using (2.4) and the pointwise ergodic theorem. \square

Bourgain-Furstenberg-Katznelson-Ornstein criterion. Our main tool for proving convergence u.a.e. is the following criterion that reduces the search for a universal set of $x \in X$ (that a priori involves uncountably many systems Y) to a problem about X^2 .

Proposition 2.11 ([BFKO89, Proposition]). Let (X, T) be an ergodic measure-preserving system and $f \in L^\infty(X) \cap \mathcal{Z}_1(X)^\perp$. Assume that $x \in X$ is generic for f and

$$\frac{1}{N} \sum_{n=1}^N f(T^n x) f(T^n \xi) \rightarrow 0 \quad \text{for a.e. } \xi \in X.$$

Then for every measure-preserving system (Y, S) and $g \in L^\infty(Y)$ we have

$$\frac{1}{N} \sum_{n=1}^N f(T^n x) g(S^n y) \rightarrow 0 \quad \text{for a.e. } y \in Y.$$

This has been generalized to averages along Følner sequences in countable amenable groups satisfying the Tempelman condition [OW92, §3]. It seems plausible that a corresponding statement for tempered Følner sequences (for which the pointwise ergodic theorem is known to hold by [Lin01]) would also be true, but we concentrate on the standard Cesàro averages.

A measure-theoretic lemma. The next lemma is our main tool for dealing with cube measures. Informally, it shows that a certain kind of universality for $\mu^{[1]} \otimes \nu^{[1]}$ implies some universality for $(\mu \times \nu)^{[1]}$.

Recall that, for ergodic measure-preserving systems $(X, \mu), (Y, \nu)$, the projection onto the invariant factor of $X \times Y$ has the form $\phi(x, y) = \psi(\pi_1(x), \pi_1(y))$, where π_1 are projections onto the Kronecker factors and ψ is the quotient map of $\mathcal{Z}_1(X) \times \mathcal{Z}_1(Y)$ by the orbit closure of the identity. To see this, recall that by Lemma 2.7 the function $f \otimes g$, $f \in L^\infty(X)$, $g \in L^\infty(Y)$, is orthogonal to the invariant factor of $X \times Y$ whenever $f \perp \mathcal{Z}_1(X)$ or $g \perp \mathcal{Z}_1(Y)$. Thus the invariant sub- σ -algebra on $X \times Y$ is contained in $\mathcal{Z}_1(X) \times \mathcal{Z}_1(Y)$, i.e. it is (isomorphic to) the invariant sub- σ -algebra of a product of two compact group rotations (cf. e.g. [Rud95, Theorem 1.9]). In particular, for an ergodic system Y the invariant factor of $Y \times Y$ is isomorphic to $\mathcal{Z}_1(Y)$.

Lemma 2.12. *Let $(X, \mu), (Y, \nu)$ be ergodic measure-preserving systems and fix measure disintegrations*

$$\mu = \int_{\kappa \in \mathcal{Z}_1(X)} \mu_\kappa d\kappa, \quad \nu = \int_{\lambda \in \mathcal{Z}_1(Y)} \mu_\lambda d\lambda.$$

This induces an ergodic decomposition

$$\nu \otimes \nu = \int_{\lambda \in \mathcal{Z}_1(Y)} (\nu \otimes \nu)_\lambda d\lambda, \quad (\nu \otimes \nu)_\lambda = \int_{\lambda' \in \mathcal{Z}_1(Y)} \nu_{\lambda'} \otimes \nu_{\lambda'\lambda^{-1}} d\lambda'.$$

Let $x \in X$ and $\Lambda \subset \mathcal{Z}_1(Y)$ be a full measure set. Assume that for μ -a.e. ξ and every $\lambda \in \Lambda$, for $(\nu \otimes \nu)_\lambda$ -a.e. (η, η') one some statement $P(x, \xi, \eta, \eta')$ holds. Then $P(x, \xi, y, \eta)$ also holds for ν -a.e. y and $\tilde{m}_{x,y}$ -a.e. (ξ, η) , where

$$\tilde{m}_{x,y} = \int_{\kappa \in \mathcal{Z}_1(X), \lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x), \pi_1(y)) = \psi(\kappa, \lambda)} \mu_\kappa \otimes \nu_\lambda d(\kappa, \lambda),$$

the homomorphism ψ is as above and the integral is taken over an affine subgroup with respect to its Haar measure.

Proof. Recall that $\ker \psi$ has full projections on both coordinates. Therefore, for every x there is a full measure set of ξ such that the set Λ has full measure in $\{\lambda : \psi(\pi_1(x)\pi_1(\xi)^{-1}, \lambda) = \text{id}\}$ (note that this is a closed affine subgroup of $\mathcal{Z}_1(Y)$ that therefore has a Haar measure).

In particular, for a full measure set of ξ (that depends on Y) the hypothesis holds for a.e. λ with $\psi(\pi_1(x)\pi_1(\xi)^{-1}, \lambda) = \text{id}$, i.e. we have $P(x, \cdot)$ for a set of full measure w.r.t. the measure

$$\begin{aligned} & \int_{\xi \in X} \delta_\xi \otimes \int_{\lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x)\pi_1(\xi)^{-1}, \lambda) = \text{id}} (\nu \otimes \nu)_\lambda d\lambda d\mu(\xi) \\ &= \int_{\kappa \in \mathcal{Z}_1(X)} \int_{\lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x)\kappa^{-1}, \lambda) = \text{id}} \mu_\kappa \otimes (\nu \otimes \nu)_\lambda d\lambda d\kappa \\ &= \int_{\kappa \in \mathcal{Z}_1(X)} \int_{\lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x)\kappa^{-1}, \lambda) = \text{id}} \mu_\kappa \otimes \int_{\lambda' \in \mathcal{Z}_1(Y)} \nu_{\lambda'} \otimes \nu_{\lambda'\lambda^{-1}} d\lambda' d\lambda d\kappa \\ &= \int_{\kappa \in \mathcal{Z}_1(X)} \int_{\lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x)\kappa^{-1}, \lambda) = \text{id}} \mu_\kappa \otimes \int_{y \in Y} \delta_y \otimes \nu_{\pi(y)\lambda^{-1}} d\nu(y) d\lambda d\kappa \\ &= \int_{y \in Y} \int_{\kappa \in \mathcal{Z}_1(X)} \int_{\lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x)\kappa^{-1}, \lambda) = \text{id}} \mu_\kappa \otimes \delta_y \otimes \nu_{\pi(y)\lambda^{-1}} d\lambda d\kappa d\nu(y) \end{aligned}$$

$$\begin{aligned}
&= \int_{y \in Y} \int_{\kappa \in \mathcal{X}_1(X), \lambda \in \mathcal{X}_1(Y) : \psi(\pi_1(x), \pi_1(y)) = \psi(\kappa, \lambda)} \mu_\kappa \otimes \delta_y \otimes \nu_\lambda d(\kappa, \lambda) dv(y) \\
&= \int_{y \in Y} \delta_y \otimes \tilde{m}_{x,y} dv(y).
\end{aligned}$$

This gives $P(x, \xi, y, \eta)$ for ν -a.e. y and $\tilde{m}_{x,y}$ -a.e. pair (ξ, η) as required. \square

The next lemma provides us with means for using the measure $\tilde{m}_{x,y}$ in a higher step setting.

Lemma 2.13. *Let $(Z, g), (Z', g')$ be ergodic nilrotations and $\psi : \mathcal{X}_1(Z) \times \mathcal{X}_1(Z') \rightarrow H$ the factor map modulo the orbit closure of $(\pi_1(g), \pi_1(g'))$. Then for every $\lambda \in \mathcal{X}_1(Z)$ and a.e. $\lambda' \in \mathcal{X}_1(Z')$ the rotation by (g, g') on the nilmanifold*

$$N_{\lambda, \lambda'} = \{(z, z') \in Z \times Z' : \psi(\pi_1(z), \pi_1(z')) = \psi(\lambda, \lambda')\}$$

is uniquely ergodic.

Proof. By [Lei05, 2.17-2.20] it suffices to prove ergodicity to obtain unique ergodicity.

Since $N_{\lambda, \lambda'}$ only depends on $\psi(\lambda, \lambda')$ and $\ker \psi$ has full projection on $\mathcal{X}_1(Z)$ it suffices to verify the conclusion for a full measure set of (λ, λ') . For this end it suffices to check that for any $f \in C(Z), f' \in C(Z')$ the limit of the ergodic averages of $f \otimes f'$ is essentially constant on $N_{\lambda, \lambda'}$. We decompose $f = f_\perp + f_{\mathcal{X}}$ with $f_\perp \perp \mathcal{X}_1(Z)$ and $f_{\mathcal{X}} \in L^\infty(\mathcal{X}_1(Z))$, and analogously for f' . For $f_{\mathcal{X}} \otimes f'_{\mathcal{X}}$ the limit is essentially constant on $N_{\lambda, \lambda'}$ for any (λ, λ') since the rotation is ergodic on $(\pi_1 \times \pi_1)(N_{\lambda, \lambda'})$.

On the other hand, the limit of the ergodic averages of tensor products involving f_\perp vanishes on $Z \times Z'$ a.e. by Lemma 2.7, hence also a.e. on a.e. fiber $N_{\lambda, \lambda'}$. \square

3. RETURN TIMES THEOREM ON CUBE SPACES

In order to concisely state our central result, Theorem 3.2, we need a cube version of Definition 1.1. Recall that we write $f_i^{[l]} = \otimes_{\epsilon \in \{0,1\}^l} f_{i,\epsilon}$, where $f_{i,\epsilon} \in L^\infty(X_i)$.

Definition 3.1. Let P be a statement about ergodic regular measure-preserving systems (X_i, μ_i, T_i) , functions $f_i^{[l]}$ and points $x_i \in X_i^{[l]}$, $i = 0, \dots, k$. We say that P holds for $[l]$ -universally almost every ($[l]$ -u.a.e.) tuple x_0, \dots, x_k if

- (0) For every system (X_0, μ_0, T_0, D_0) there exists a measurable set $\tilde{X}_0^{[l]} \subset X_0^{[l]}$ such that for every $y_0 \in Y_{0,l}$ we have $m_{y_0}(\tilde{X}_0^{[l]}) = 1$ and
- (1) for every system (X_1, μ_1, T_1, D_1) there exists a measurable set $\tilde{X}_1^{[l]} \subset X_0^{[l]} \times X_1^{[l]}$ such that for every $\vec{x}_0 \in \tilde{X}_0^{[l]}$ and every $y_1 \in Y_{1,l}$ we have $m_{y_1}\{x_1 : (\vec{x}_0, x_1) \in \tilde{X}_1^{[l]}\} = 1$ and
- \vdots
- (k) for every system (X_k, μ_k, T_k, D_k) there exists a measurable set $\tilde{X}_k^{[l]} \subset X_0^{[l]} \times \dots \times X_k^{[l]}$ such that for every $\vec{x}_{k-1} \in \tilde{X}_{k-1}^{[l]}$ and every $y_k \in Y_{k,l}$ we have $m_{y_k}\{x_k : (\vec{x}_{k-1}, x_k) \in \tilde{X}_k^{[l]}\} = 1$ and

we have $P(f_0^{[l]}, \dots, f_k^{[l]}, \vec{x}_k)$ for every $\vec{x}_k \in \tilde{X}_k$ and any $f_{i,\epsilon} \in D_i$, $0 \leq i \leq k$, $\epsilon \in \{0,1\}^l$.

With this definition “u.a.e.” corresponds to “[0]-u.a.e.”.

Our main theorem below states that certain nilfactors are characteristic for return time averages on cube spaces.

Theorem 3.2. *For any $k, l \in \mathbb{N}$ the ergodic averages of $\otimes_{i=0}^k f_i^{[l]}$ converge $[l]$ -u.a.e. If in addition*

$$\text{CF}(k, l) \quad \exists \epsilon \in \{0,1\}^l \text{ s.t. } f_{0,\epsilon} \perp \mathcal{X}_{k+l}(X_0) \text{ or } f_{i,\epsilon} \perp \mathcal{X}_{k+l+1-i}(X_i) \text{ for some } 1 \leq i \leq k$$

then the limit vanishes $[l]$ -u.a.e.

The case $l = 0$ of Theorem 3.2 is RTT(k) with additional information CF(k,0) about characteristic factors.

The proof is by induction on k and the inductive loop of the proof spans this whole section. The base case $k = 0$ follows by definition of $Y_{0,l}$ and the pointwise ergodic theorem.

Thus we assume Theorem 3.2 for some fixed $k \in \mathbb{N}$ and prove all intermediate results for this value of k assuming that they hold for lower values of k . The base cases of the intermediate results will be proved as we state them.

In order to pass from k to $k + 1$ in Theorem 3.2 we write

$$(3.3) \quad X_0^{[l+1]} \times \cdots \times X_k^{[l+1]} = (X_0^{[l]} \times \cdots \times X_k^{[l]})^2 =: X^2.$$

Our aim is to apply Proposition 2.11 with this X and $Y = X_{k+1}^{[l]}$. The remaining part of this section is devoted to verification of the hypotheses of that proposition. This involves the following steps. First we use RTT(k) to construct a certain universal measure disintegration with built-in genericity properties on a product of ergodic systems (Theorem 3.4). We use characteristic factors for RTT(k) to represent measures in this disintegration in a different way. Finally, we verify a certain instance of RTT(k+1) (Lemma 3.12).

Universal disintegration of product measures. The return times theorem can be seen as a statement about measure disintegration, cf. [ALR95, Theorem 4] for the special case $k = 1$.

Theorem 3.4. *Let (X_i, μ_i, T_i, D_i) , $i = 0, \dots, k$, be systems. Then $[l]$ -u.a.e. x_0, \dots, x_k is generic for some measure m_{x_0, \dots, x_k} on $X_0^{[l]} \times \cdots \times X_k^{[l]}$ and every function $\otimes_{i=0}^k f_i^{[l]}$, $f_{i,\epsilon} \in D_i$.*

Moreover, for $[l]$ -u.a.e. x_0, \dots, x_{k-1} and every $y_k \in Y_{l,k}$ one has

$$(3.5) \quad m_{x_0, \dots, x_{k-1}} \otimes m_{y_k} = \int m_{x_0, \dots, x_k} dm_{y_k}(x_k).$$

Proof. By Theorem 3.2 with $l = 0$ we obtain convergence of the averages

$$\frac{1}{N} \sum_{n=1}^N \prod_{i=0}^k f_i^{[l]}(T_i^n x_i)$$

for $[l]$ -u.a.e. x_0, \dots, x_k and any $f_{i,\epsilon} \in D_i$. For continuous functions $f_{i,\epsilon} \in D_i$ we define $m_{x_0, \dots, x_k}(\otimes_{i=0}^k f_i^{[l]})$ as the limit of these averages. By the Stone-Weierstraß theorem these tensor products span a dense subspace $C(X_0^{[l]} \times \cdots \times X_k^{[l]})$, so by density the above (bounded) linear form admits a unique continuous extension.

In order to obtain (3.5) it suffices to verify that the integrals of functions of the form $\otimes_{i=0}^k f_i^{[l]}$, $f_{i,\epsilon} \in D_i$, with respect to both measures coincide. By genericity and the dominated convergence theorem we have for $[l]$ -u.a.e. x_0, \dots, x_{k-1} that

$$\begin{aligned} & \int \int \otimes_{i < k} f_i^{[l]} \otimes f_k^{[l]} dm_{x_0, \dots, x_k} dm_{y_k}(x_k) \\ &= \int \lim_N \frac{1}{N} \sum_{n=1}^N \prod_{i < k} f_i^{[l]}(T_i^n x_i) \cdot f_k^{[l]}(T_k^n x_k) dm_{y_k}(x_k) \\ &= \lim_N \frac{1}{N} \sum_{n=1}^N \prod_{i < k} f_i^{[l]}(T_i^n x_i) \cdot \int f_k^{[l]}(T_k^n x_k) dm_{y_k}(x_k) \\ &= \int \otimes_{i < k} f_i^{[l]} dm_{x_0, \dots, x_{k-1}} \int f_k^{[l]} dm_{y_k} \end{aligned}$$

as required. \square

Properties of the universal disintegration. We will now represent the measure m_{x_0, \dots, x_k} for $[l]$ -u.a.e. x_0, \dots, x_k in the form $\tilde{m}_{x,y}$ in the notation of Lemma 2.12. At this step we have to use the information about characterisitic factors. We begin with a preliminary observation.

Lemma 3.6. *If some property P holds $[l]$ -u.a.e. then, for $[l]$ -u.a.e. x_0, \dots, x_k , P holds m_{x_0, \dots, x_k} -a.e.*

Proof. For $k = 0$ this follows from (3.5). Assume that the conclusion is known for $k - 1$ and show it for k .

By the induction hypothesis, for $[l]$ -u.a.e. x_0, \dots, x_{k-1} , $m_{x_0, \dots, x_{k-1}}$ -a.e., for every $y_k \in Y_{k,l}$, P holds m_{y_k} -a.e. in x_k . The conclusion follows from (3.5). \square

Lemma 3.7. *For $[l]$ -u.a.e. x_0, \dots, x_k we have*

$$m_{x_0, \dots, x_k} = \tilde{m}_{x,y},$$

where we use the notation of Lemma 2.12 with $(X, \mu) = (X_0^{[l]} \times \dots \times X_{k-1}^{[l]}, m_{x_0, \dots, x_{k-1}})$, $(Y, \nu) = (X_k^{[l]}, m_{x_k})$, $x = (x_0, \dots, x_{k-1})$ and $y = x_k$.

Proof. To verify that the measures coincide it suffices to check that the integrals of functions of the form $\otimes_{i=0}^k f_i^{[l]}$, $f_{i,\epsilon} \in D_i$ coincide. For this end consider the splittings $f_{i,\epsilon} = f_{i,\epsilon,\perp} + f_{i,\epsilon,\mathcal{Z},j} + f_{i,\epsilon,err,j}$, $j \in \mathbb{N}$, given by Dec($k + l + 1 - i$).

Projections of tensor products that involve $f_{i,\epsilon,\perp}$ on one of the Kronecker factors vanish a.e. for $[l]$ -u.a.e. x, y by Corollary 3.13 for $k - 1$ that is part of the induction hypothesis for this section. Since $\ker \psi$ has full projections on both coordinates the corresponding integrals w.r.t. $\tilde{m}_{x,y}$ also vanish. The integrals w.r.t. $m_{x,y}$ vanish for $[l]$ -u.a.e. x, y by Theorem 3.2.

For the main terms we have

$$\begin{aligned} (3.8) \quad & \int \otimes_{i=0}^k f_{i,\mathcal{Z},j}^{[l]} d\tilde{m}_{x,y} \\ &= \int_{\kappa \in \mathcal{Z}_1(X), \lambda \in \mathcal{Z}_1(Y): \psi(\pi_1(x), \pi_1(y)) = \psi(\kappa, \lambda)} \mathbb{E}(\otimes_{i=0}^{k-1} f_{i,\mathcal{Z},j}^{[l]} | \mathcal{Z}_1(X))(\kappa) \mathbb{E}(f_{k,\mathcal{Z},j}^{[l]} | \mathcal{Z}_1(Y))(\lambda) d(\kappa, \lambda). \end{aligned}$$

Since the underlying nilmanifold of a nilrotation is a bundle of nilmanifolds over its Kronecker factor, the conditional expectation above is just integration in the fibers, and by uniqueness of the Haar measure the whole integral equals

$$\int_{\kappa \in Z_j, \lambda \in Z'_j: \psi(\pi_1(x), \pi_1(y)) = \psi(\pi_1(\kappa), \pi_1(\lambda))} \otimes_{i=0}^{k-1} f_{i,\mathcal{Z},j}^{[l]}(\kappa) f_{k,\mathcal{Z},j}^{[l]}(\lambda) d(\kappa, \lambda),$$

where Z_j is the orbit closure of x in $\prod_{i=0}^{k-1} Z_{i,j}^{[l]}$ and Z'_j is the orbit closure of y in $Z_{k,j}^{[l]}$. By Lemma 2.13, the above fibers of $Z_j \times Z'_j$ are uniquely ergodic for every x and a.e. y , and the integral then equals

$$\lim_N \frac{1}{N} \sum_{n=1}^N \otimes_{i=0}^{k-1} f_{i,\mathcal{Z},j}^{[l]}(T^n x) f_{k,\mathcal{Z},j}^{[l]}(S^n y) = \int \otimes_{i=0}^k f_{i,\mathcal{Z},j}^{[l]} d\tilde{m}_{x,y}.$$

It remains to treat the error terms, i.e. the case $f_{i',\epsilon'} = f_{i',\epsilon',err,j}$ for some i', ϵ' . By Lemma 2.9(3), for $[l]$ -u.a.e. x, y we have

$$\int \otimes_{i=0}^k f_i^{[l]} d\tilde{m}_{x,y} \lesssim \|f_{i',\epsilon'}\|_{L^1(\mu_i)} \rightarrow 0 \quad \text{as } j \rightarrow \infty.$$

Similarly, we have $\int |\otimes_{i=0}^{k-1} f_i^{[l]}| d\mathbf{m}_x \lesssim \|f_{i',\epsilon'}\|_{L^1(\mu_i)}$ if $i' < k$ and $\int |f_k^{[l]}| d\mathbf{m}_{y_k} \lesssim \|f_{k,\epsilon}\|_{L^1(\mu_i)}$ if $i' = k$ for $[l]$ -u.a.e. x, y . This implies that either $\mathbb{E}(\otimes_{i=0}^{k-1} f_i^{[l]} | \mathcal{Z}_1(X))$ or $\mathbb{E}(f_k^{[l]} | \mathcal{Z}_1(Y))$ converges to zero in probability for $[l]$ -u.a.e. x, y , so

$$\int \otimes_{i=0}^k f_i^{[l]} d\tilde{\mathbf{m}}_{x,y} \rightarrow 0 \quad \text{as } j \rightarrow \infty$$

for $[l]$ -u.a.e. x_0, \dots, x_k since $\ker \psi$ has full projections on coordinates. \square

Corollary 3.9. *For $[l]$ -u.a.e. x_0, \dots, x_k the measure $\mathbf{m}_{x_0, \dots, x_k} = \tilde{\mathbf{m}}_{x,y}$ is ergodic.*

Note that even for a non-ergodic invariant measure on a regular system there may exist generic points, so the mere fact that \vec{x} is generic for $\mathbf{m}_{\vec{x}}$ does not suffice.

Proof. In order to see that $\mathbf{m}_{x_0, \dots, x_k}$ is ergodic it suffices to verify that for any continuous functions $f_{i,\epsilon} \in C(X_i)$ we have

$$(3.10) \quad \lim_N \frac{1}{N} \sum_{n=1}^N \otimes_{i=0}^k f_i^{[l]}(T^n \vec{\xi}) = \int \otimes_{i=0}^k f_i^{[l]} d\mathbf{m}_{x_0, \dots, x_k} \quad \text{for } \mathbf{m}_{x_0, \dots, x_k}\text{-a.e. } \vec{\xi}.$$

Recall that for $[l]$ -u.a.e. x_0, \dots, x_k the limit on the left-hand side of (3.10) exists for $\mathbf{m}_{x_0, \dots, x_k}$ -a.e. $\vec{\xi}$ by Lemma 3.6 and equals $\int \otimes_{i=0}^k f_i^{[l]} d\mathbf{m}_{\vec{\xi}}$. Splitting the $f_{i,\epsilon}$'s as before it suffices to verify (3.10) for the main terms, and this follows directly from (3.8). \square

The sufficient special case of convergence to zero. The last hypothesis of Proposition 2.11 is a certain special case of its conclusion. Recall that we already have u.a.e. convergence to zero on X^2 (as defined in (3.3)), but not yet in the required sense. This is now corrected using Lemma 2.12.

Lemma 3.11 (Change of order in the cube construction). *Let $l \in \mathbb{N}$ and P be a statement about points of $\prod_{i=0}^k X_i^{[l+1]}$. Assume that for $[l+1]$ -u.a.e. x_0, \dots, x_k we have $P(x_0, \dots, x_k)$.*

Then for $[l]$ -u.a.e. x_0, \dots, x_k , for $\mathbf{m}_{x_0, \dots, x_k}$ -a.e. x' , we have $P(x_0, \dots, x_k, x')$.

Strictly speaking, the coordinates of x' in (x_0, \dots, x_k, x') should be attached to x_0, \dots, x_k but we do not want to introduce additional notation at this point.

Proof. The base case $k = 0$ follows directly from (2.10).

Assume now that $k > 0$. By the inductive hypothesis of this section the conclusion holds for $k-1$, so for $[l]$ -u.a.e. x_0, \dots, x_{k-1} , for $\mathbf{m}_{x_0, \dots, x_{k-1}}$ -a.e. x' , for every $y_k \in Y_{k,l+1}$ and \mathbf{m}_{y_k} -a.e. x_k , we have $P(x_0, \dots, x_{k-1}, x', x_k)$.

Using (2.10) we can rewrite the emphasized part of the statement as “for $\mathbf{m}_{x_0, \dots, x_{k-1}}$ -a.e. x' , for every $\tilde{y}_k \in Y_{k,l}$, for every ergodic component μ_e of $(\mathbf{m}_{\tilde{y}_k})^2$ from a fixed full measure set, for μ_e -a.e. x_k ” The conclusion follows by Lemma 2.12 and Lemma 3.7. \square

Lemma 3.12. *Let $l, l' \in \mathbb{N}$ and assume $CF(k, l + l')$. Then for $[l]$ -u.a.e. $\vec{x}_0 = (x_0, \dots, x_k)$, for $\mathbf{m}_{\vec{x}_0}$ -a.e. \vec{x}_1, \dots , for $\mathbf{m}_{\vec{x}_0, \dots, \vec{x}_{l'-1}}$ -a.e. $\vec{x}_{l'}$, the ergodic averages of the function $\otimes_{i=0}^k f_i^{[l+l']}$ converge to zero at $(\vec{x}_0, \dots, \vec{x}_{l'})$.*

Again, the tensor product $\otimes_{i=0}^k f_i^{[l+l']}$ should be arranged in a different order, but in our opinion the above notation makes our goal more clear: it is not the function but the order in which we build the product space that changes.

Proof. We use induction on l' . The case $l' = 0$ is precisely Theorem 3.2. Assume that the conclusion is known for $l+1$ and $l' - 1$. The claim for l and l' follows by Lemma 3.11. \square

Corollary 3.13. *Let $l, l' \in \mathbb{N}$ and assume $CF(k, l + l')$. Then for $[l]$ -u.a.e. x_0, \dots, x_k we have $f_0^{[l]} \otimes \dots \otimes f_k^{[l]} \perp \mathcal{Z}_{l'}(\mathbf{m}_{x_0, \dots, x_k})$.*

Proof. This follows from Lemma 3.12 by Lemma 3.6, the definition of cube measures (2.3), the characterization of uniformity seminorms (2.5) and the ergodic theorem. \square

Proof of Theorem 3.2 for $k + 1$. Assume first $\text{CF}(k, l)$. Then Lemma 3.12 with $l' = 1$ states that for $[l]$ -u.a.e. $x = (x_0, \dots, x_k)$, for $\mathbf{m}_{x_0, \dots, x_k}$ -a.e. x' , for any $f_{i, \epsilon} \in D_i$ we have

$$\lim_N \frac{1}{N} \sum_{n=1}^N \otimes_{i=0}^k f_i^{[l]}((\otimes_{i=0}^k T_i^{[l]})^n x) \cdot \otimes_{i=0}^k f_i^{[l]}((\otimes_{i=0}^k T_i^{[l]})^n x') = 0.$$

For $[l]$ -u.a.e. x_0, \dots, x_k we obtain genericity w.r.t. $\mathbf{m}_{x_0, \dots, x_k}$ by Theorem 3.4, ergodicity of $\mathbf{m}_{x_0, \dots, x_k}$ by Corollary 3.9 and orthogonality of $\otimes_{i=0}^k f_i^{[l]}$ to the Kronecker factor of $\mathbf{m}_{x_0, \dots, x_k}$ by Corollary 3.13, so Proposition 2.11 with $X = (X_0^{[l]} \times \dots \times X_k^{[l]}, \mathbf{m}_{x_0, \dots, x_k})$ and $Y = (X_{k+1}^{[l]}, \mathbf{m}_{y_{k+1}})$ implies the claimed convergence to zero $[l]$ -u.a.e.

This takes care of the terms $f_{i, \epsilon, \perp}$ in the splittings $f_{i, \epsilon} = f_{i, \epsilon, \perp} + f_{i, \epsilon, \mathcal{Z}, j} + f_{i, \epsilon, \text{err}, j}$ given by $\text{Dec}(k + l + 1 - i)$ (resp. $k + l$ for $i = 0$). By an approximation argument like in the proof of Lemma 3.7 it suffices to consider the main terms, so we may assume that $\prod_{i=0}^k f_i^{[l]}((T_i^{[l]})^n x)$ is a nilsequence. The claimed convergence a.e. in x_{k+1} then follows from the Wiener-Wintner theorem for nilsequences [HK09, Theorem 2.22]. The limit is zero a.e. in x_{k+1} provided that $f_{k+1, \epsilon} \perp \mathcal{Z}_{l+1}(X_{k+1})$ for some ϵ by definition of $Y_{k+1, l}$ and Lemma 2.7. \square

4. WIENER-WINTNER RETURN TIMES THEOREM FOR NILSEQUENCES

The first step in the proof is the identification of characteristic factors in the spirit of [ALR95, §4]. We refer to [EZK12] for the notation.

Lemma 4.1. *Let $f_i \in L^\infty(X_i)$, $i = 0, \dots, k$, and assume $\text{CF}(k, l)$. Let further G/Γ be a nilmanifold, G_\bullet a filtration on G of length l , and let a Mal'cev basis adapted to G_\bullet be fixed. Then for u.a.e. x_0, \dots, x_k we have*

$$(4.2) \quad \lim_{N \rightarrow \infty} \sup_{g \in P(\mathbb{Z}, G_\bullet), F \in W^{\tilde{k}, 2^l}(G/\Gamma)} \|F\|_{W^{\tilde{k}, 2^l}(G/\Gamma)}^{-1} \left| \frac{1}{N} \sum_{n=1}^N F(g(n)\Gamma) \prod_{i=0}^k f_i(T_i^n x_i) \right| = 0,$$

where $\tilde{k} = \sum_{r=1}^l (d_r - d_{r+1}) \binom{l}{r-1}$.

Proof. By Corollary 3.13 we have $\otimes_{i=0}^k f_i \perp \mathcal{Z}_l(\mathbf{m}_{\vec{x}})$ for u.a.e. $\vec{x} \in X_0 \times \dots \times X_k$ and by Theorem 3.4 u.a.e. \vec{x} is fully generic for $\otimes_i f_i$ w.r.t. $\mathbf{m}_{\vec{x}}$. The claim follows by the uniform Wiener-Wintner theorem [EZK12, Theorem 4.1]. \square

Theorem 1.3 now follows from equidistribution results on nilmanifolds.

Proof of Theorem 1.3 for the given k . By Lemma 4.1 it suffices to consider $f_i \in L^\infty(\mathcal{Z}_{l+k+1-i}(X_i))$. By the pointwise ergodic theorem we can assume that each f_i is a smooth function on a nilsystem factor of X_i . The conclusion now follows from equidistribution results on nilmanifolds [Lei05, Theorem B]. \square

REFERENCES

- [ALR95] I. Assani, E. Lesigne, and D. Rudolph, *Wiener-Wintner return-times ergodic theorem*, Israel J. Math. **92** (1995), no. 1-3, 375–395. MR1357765 (97e:28011) ↑2, 7, 10
- [AP12a] I. Assani and K. Presser, *A Survey of the Return Times Theorem*, ArXiv e-prints (September 2012), available at 1209.0856. ↑2
- [AP12b] I. Assani and K. Presser, *Pointwise characteristic factors for the multiterm return times theorem*, Ergodic Theory Dynam. Systems **32** (2012), no. 2, 341–360. MR2901351 ↑2
- [BFKO89] J. Bourgain, H. Furstenberg, Y. Katznelson, and D. S. Ornstein, *Appendix on return-time sequences*, Inst. Hautes Études Sci. Publ. Math. **69** (1989), 42–45. MR1557098 ↑1, 4

- [Bir31] G. D. Birkhoff, *Proof of the ergodic theorem.*, Proc. Natl. Acad. Sci. USA **17** (1931), 656–660. ↑2
- [EZK12] T. Eisner and P. Zorin-Kranich, *Uniformity in the Wiener-Wintner theorem for nilsequences*, 2012. Preprint, arXiv:1208.3977. ↑10
- [HK05] B. Host and B. Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2) **161** (2005), no. 1, 397–488. MR2150389 (2007b:37004) ↑3
- [HK09] ———, *Uniformity seminorms on ℓ^∞ and applications*, J. Anal. Math. **108** (2009), 219–276. MR2544760 (2010j:11018) ↑3, 10
- [HKM10] B. Host, B. Kra, and A. Maass, *Nilsequences and a structure theorem for topological dynamical systems*, Adv. Math. **224** (2010), no. 1, 103–129. MR2600993 (2011h:37014) ↑3
- [Lei05] A. Leibman, *Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold*, Ergodic Theory Dynam. Systems **25** (2005), no. 1, 201–213, available at <http://www.math.osu.edu/~leibman.1/preprints/PolNilRtn.pdf>. MR2122919 (2006j:37004) ↑6, 10
- [Lin01] E. Lindenstrauss, *Pointwise theorems for amenable groups*, Invent. Math. **146** (2001), no. 2, 259–295. MR1865397 (2002h:37005) ↑4
- [OW92] D. Ornstein and B. Weiss, *Subsequence ergodic theorems for amenable groups*, Israel J. Math. **79** (1992), no. 1, 113–127. MR1195256 (94g:28024) ↑4
- [Rud95] D. J. Rudolph, *Eigenfunctions of $T \times S$ and the Conze-Lesigne algebra*, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), 1995, pp. 369–432. MR1325712 (96k:28025) ↑5
- [Rud98] ———, *Fully generic sequences and a multiple-term return-times theorem*, Invent. Math. **131** (1998), no. 1, 199–228. MR1489899 (99c:28055) ↑1

KORTEWEG-DE VRIES INSTITUTE FOR MATHEMATICS, UNIVERSITY OF AMSTERDAM, P.O. Box 94248, 1090 GE AMSTERDAM, THE NETHERLANDS

E-mail address: zorin-kranich@uva.nl

URL: <http://staff.science.uva.nl/~pavelz/>